What is claimed is:

1. A method of detecting a person from an image provided by a digital video camera, the method comprising:

(a) initializing parameters used for a user detection;

(b) determining a current mode;

(c) when it is determined that the current mode is the face detection mode, performing a face detection on each of a plurality of image frames, and when a face is not detected even after a predetermined number of face detection trials, selecting a motion detection mode and going back to step (b); and

(d) when it is determined that the current mode is the motion detection mode, performing a motion detection on a predetermined number of image frames at a time, and when at least a predetermined number of motion detections are performed successfully within a predetermined period of time, going back to step (a).

2. The method of claim 1, wherein step (c) comprises detecting a face candidate region using a Gabor wavelet transformation.

3. The method of claim 2, wherein step (c) comprises detecting a face using a low-resolution support vector machine (SVM) and a high resolution SVM.

4. The method of claim 3, wherein step (c) comprises:

(c1) receiving a single image frame;

(c2) increasing the number of face detection trials by 1;

(c3) detecting the face candidate region by training a plurality of model face figures based on an M-grid Gabor wavelet transformation and determining whether a face candidate region has been detected;

(c4) when it is determined that a face candidate region has been detected, performing a face detection using the low-resolution SVM and determining whether a face has been detected;

(c5) when it is determined that a face has been detected using the low-resolution SVM, performing a face detection using the high-resolution SVM and determining whether a face has been detected;

(c6) when it is determined that a face has been detected using the high-resolution SVM, storing the current image frame, initializing the number of face detection trials, and going back to step (c1);

(c7) when it is determined that a face candidate region has not been detected in step (c3) or when it is determined that a face has not been detected in step (c4) or (c5), determining whether the number of face detection trials is equal to or greater than a first threshold value and going back step (c1) when it is determined that the number of face detection trials is less than the first threshold value; and

(c8) when it is determined that the number of face detection trials is equal to or greater than the first threshold value, selecting the motion detection mode and going back to step (b).

5. The method of claim 4, wherein in step (c3), M-grid intervals for training the model face figures are determined based on an inter ocular distance and a distance between a mouth and a line connecting two eyes.

6. The method of claim 4, wherein steps (c4) and (c5) comprise performing face detection using a classification based on an SVM trained in principal component analysis (PCA) subspaces.

7. The method of claim 6, wherein step (c4) comprises selecting an area having a predetermined size on the basis of a predetermined position around the face candidate region detected in step (c3), performing a face detection on the selected area using the low-resolution SVM, and repeating the face detection while extending the selected area so that a predetermined area including the face candidate region is fully scanned.

8. The method of claim 6, wherein step (c5) comprises selecting an area having a predetermined size on the basis of a predetermined position around the face detected in step (c4), performing a face detection on the selected area using the high-resolution SVM, and repeating the face detection while extending the selected area so that a predetermined area including the face is fully scanned.

9. The method of claim 1, wherein step (d) comprises performing the motion detection using a temporal edge detection algorithm.

10. The method of claim 9, wherein step (d) comprises:

(d1) receiving a predetermined number of image frames;

(d2) increasing the number of motion detection trials by 1;

(d3) detecting a motion in the image frames using the temporal edge detection algorithm and variance of pixel values with respect to time;

(d4) determining whether a motion has been detected;

(d5) when it is determined that a motion has been detected, increasing the number of motion detection successes by 1 and storing the image;

(d6) determining whether conditions that the number of motion detection trials is less than a third threshold value and the number of motion detection successes is equal to or greater than a second threshold value are satisfied, and going back to step (a) when it is determined that the conditions are satisfied;

(d7) when it is determined that a motion has not been detected in step (d4) or when it is determined that the conditions are not satisfied in step (d6), determining whether the number of motion detection trials is equal to or greater than the third threshold value, and going back to step (d1) when it is determined that the number of motion detection trials is less than the third threshold value; and

(d8) when it is determined that the number of motion detection trials is equal to or greater than the third threshold value, initializing the number of motion detection trials and the number of motion detection successes and going back to step (d1).

11. The method of claim 10, wherein step (d3) comprises:

(d30) detecting pixels by finding zero crossings of a Laplacian with respect to time using the predetermined number of image frames;

(d32) calculating the variance of pixel values with respect to time using the predetermined number of image frames; and

(d34) determining whether the variance calculated for each zero crossing pixel of the Laplacian is equal to or greater than a predetermined value, and determining the pixel as a motion pixel when it is determined that the variance is equal to or greater than the predetermined value.

12. The method of claim 11, wherein when $2m$ image frames are received and "n" pixels exist in each image frame, step (d30) comprises:

(d300) multiplying each of $f(t_1)$ through $f(t_{2m-1})$ in a group of image frames by a predetermined weight and then averaging the results of multiplication to obtain Laplacian of $f(t_m)$ with respect to time, i.e., $\nabla^2 f(t_m)$;

(d302) multiplying each of $f(t_2)$ through $f(t_{2m})$ in another group of image frames by a predetermined weight and then averaging the results of multiplication to obtain Laplacian of $f(t_{m+1})$ with respect to time, i.e., $\nabla^2 f(t_{m+1})$; and

(d304) determining a pixel of interest as a zero crossing pixel when $\nabla^2 f(t_m)$ is negative and $\nabla^2 f(t_{m+1})$ is positive or when $\nabla^2 f(t_m)$ is positive and $\nabla^2 f(t_{m+1})$ is negative,

wherein $f(t_1)$ through $f(t_{2m})$ denote pixel values of the pixel of interest under processing, and steps (d300) through (d304) are repeated "n" times.

13. The method of claim 11, wherein step (d4) comprises determining whether the number of pixels determined as motion pixels in step (d3) is at least a predetermined value and determining that a motion has been detected when it is determined that the number of motion pixels is at least the predetermined value.

14. The method of claim 11, further comprising performing a Gaussian filtering on the received image frames to smooth the image frames before step (d30).

15. The method of claim 1, wherein step (d) comprises interrupting the motion detection with a predetermined period and then performing a face detection.

16. A method of detecting a motion from an image provided by a digital video camera, the method comprising:

(e1) receiving a predetermined number of image frames;

(e2) detecting a motion in the image frames using the temporal edge detection algorithm and variance of pixel values with respect to time;

(d3) determining whether a motion has been detected; and

(d4) when it is determined that a motion has been detected, storing the image.

17. The method of claim 16, wherein step (e2) comprises:

(e20) detecting pixels by finding zero crossings of a Laplacian with respect to time using the predetermined number of image frames;

(e22) calculating the variance of pixel values with respect to time using the predetermined number of image frames; and

(e24) determining whether the variance calculated for each zero crossing pixel of the Laplacian is equal to or greater than a predetermined value, and determining the pixel as a motion pixel when it is determined that the variance is equal to or greater than the predetermined value.

18. The method of claim 17, wherein when $2m$ image frames are received and "n" pixels exist in each image frame, step (e20) comprises:

(e200) multiplying each of $f(t_1)$ through $f(t_{2m-1})$ in a group of image frames by a predetermined weight and then averaging the results of multiplication to obtain $\nabla^2 f(t_m)$;

(e202) multiplying each of $f(t_2)$ through $f(t_{2m})$ in another group of image frames by a predetermined weight and then averaging the results of multiplication to obtain $\nabla^2 f(t_{m+1})$; and

(e204) determining a pixel of interest as a zero crossing pixel when $\nabla^2 f(t_m)$ is negative and $\nabla^2 f(t_{m+1})$ is positive or when $\nabla^2 f(t_m)$ is positive and $\nabla^2 f(t_{m+1})$ is negative,

wherein $f(t_1)$ through $f(t_{2m})$ denote pixel values of the pixel of interest under processing, and steps (e200) through (e204) are repeated "n" times.

19. The method of claim 17, wherein step (e3) comprises determining whether the number of pixels determined as motion pixels in step (e2) is equal to or greater than a predetermined value and determining that a motion has been detected when it is determined that the number of motion pixels is at least the predetermined value.

20. The method of claim 17, further comprising performing a Gaussian filtering on the received image frames to smooth the image frames before step (e20).

21. An apparatus for detecting a person from an image provided by digital video camera, the apparatus comprising:

a means for initializing parameters used for a user detection;

a detection mode determination means for determining a current mode;

a face detection mode for performing a face detection on each of a plurality of image frames when it is determined that the current mode is a face detection mode and selecting a motion detection mode when a face is not detected even after a predetermined number of face detection trials; and

a motion detection means for performing a motion detection on a predetermined number of image frames at a time when it is determined that the current mode is not the face detection mode and initializing the parameters when at least a predetermined number of motion detections are performed successfully within a predetermined period of time.

22. The apparatus of claim 21, wherein the motion detection means comprises a means for detecting a face candidate region using a Gabor wavelet transformation.

23. The apparatus of claim 22, wherein the motion detection means comprises a means for detecting a face using a low-resolution support vector machine (SVM) and a high resolution SVM.

24. The apparatus of claim 23, wherein the motion detection means comprises:

a means for receiving a single image frame;

a means for increasing the number of face detection trials by 1;

a face candidate region detector means for detecting the face candidate region by training a plurality of model face figures based on an M-grid Gabor wavelet transformation and determining whether a face candidate region has been detected;

a low-resolution face detector means for performing a face detection using the low-resolution SVM when it is determined that a face candidate region has been detected and determining whether a face has been detected;

a high-resolution face detector means for performing a face detection using the high-resolution SVM when it is determined that a face has been detected using the low-resolution SVM and determining whether a face has been detected;

a means for storing the current image frame, initializing the number of face detection trials, and maintaining the face detection mode, when it is determined that a face has been detected using the high-resolution SVM;

a means for determining whether the number of face detection trials is equal to or greater than a first threshold value when it is determined that a face candidate region has not been detected or when it is determined that a face has not been detected using the low-resolution or high-resolution SVM, and maintaining the face detection mode when it is determined that the number of face detection trials is less than the first threshold value; and

a means for selecting the motion detection mode when it is determined that the number of face detection trials is equal to or greater than the first threshold value.

25. The apparatus of claim 24, wherein the face candidate region detector means comprises a means for determining M-grid intervals for training the model face figures based on an inter ocular distance and a distance between a mouth and a line connecting two eyes.

26. The apparatus of claim 24, wherein the low-resolution face detector means and the high-resolution face detector means comprise a means for performing face detection using a classification based on an SVM trained in principal component analysis (PCA) subspaces.

27. The apparatus of claim 26, wherein the low-resolution face detector means comprises a means for selecting an area having a predetermined size on the basis of a predetermined position around the face candidate region detected by the face candidate region detector means, performing a face detection on the selected area using the low-resolution SVM, and repeating the face detection while extending the selected area so that a predetermined area including the face candidate region is fully scanned.

28. The apparatus of claim 26, wherein the high-resolution face detector means comprises a means for selecting an area having a predetermined size on the basis of a predetermined position around the face detected by the low-resolution

face detector means, performing a face detection on the selected area using the high-resolution SVM, and repeating the face detection while extending the selected area so that a predetermined area including the face detected by the low-resolution face detector means is fully scanned.

29. The apparatus of claim 21, wherein the motion detection means comprises a means for performing the motion detection using a temporal edge detection algorithm.

30. The apparatus of claim 29, wherein the motion detection unit comprises:

a means for receiving a predetermined number of image frames;

a mean for increasing the number of motion detection trials by 1;

a means for detecting a motion in the image frames using the temporal edge detection algorithm and variance of pixel values with respect to time;

a means for determining whether a motion has been detected;

a means for increasing the number of motion detection successes by 1 and storing the image, when it is determined that a motion has been detected;

a means for determining whether conditions that the number of motion detection trials is less than a third threshold value and the number of motion detection successes is equal to or greater than a second threshold value are satisfied, initializing the parameters when it is determined that the conditions are satisfied, and maintaining the parameters when it is determined that the conditions are not satisfied;

a means for determining whether the number of motion detection trials is equal to or greater than the third threshold value when it is determined that a motion has not been detected or when it is determined that the conditions are not satisfied, and when it is determined that the number of motion detection trials is less than the third threshold value, maintaining the parameters, receiving a predetermined number of image frames, and performing a motion detection; and

a means for initializing the number of motion detection trials and the number of motion detection successes, receiving a predetermined number of image frames, and performing a motion detection, when it is determined that the number of motion detection trials is equal to or greater than the third threshold value.

31. The apparatus of claim 30, wherein the means for detecting a motion using the temporal edge detection algorithm and the variance comprises:

a means for detecting pixels by finding zero crossings of a Laplacian with respect to time using the predetermined number of image frames;

a means for calculating the variance of pixel values with respect to time using the predetermined number of image frames; and

a means for determining whether the variance calculated for each zero crossing pixel of the Laplacian is equal to and greater than a predetermined value, and determining the pixel as a motion pixel when it is determined that the variance is at least the predetermined value.

32. The apparatus of claim 31, wherein the means for detecting pixels by finding zero crossings of the Laplacian with respect to pixel coordinates and time comprises:

a means for multiplying each of $f(t_1)$ through $f(t_{2m-1})$ in a group of image frames by a predetermined weight and then averaging the results of multiplication to obtain Laplacian of $f(t_m)$ with respect to time, i.e., $\nabla^2 f(t_m)$;

a means for multiplying each of $f(t_2)$ through $f(t_{2m})$ in another group of image frames by a predetermined weight and then averaging the results of multiplication to obtain Laplacian of $f(t_{m+1})$ with respect to time, i.e., $\nabla^2 f(t_{m+1})$; and

a means for determining a pixel of interest as a zero crossing pixel when $\nabla^2 f(t_m)$ is negative and $\nabla^2 f(t_{m+1})$ is positive or when $\nabla^2 f(t_m)$ is positive and $\nabla^2 f(t_{m+1})$ is negative,

wherein $2m$ image frames are received, "n" pixels exist in each image frame, $f(t_1)$ through $f(t_{2m})$ denote pixel values of the pixel of interest under processing, and operations of the above three means are repeated "n" times.

33. The apparatus of claim 31, wherein the means for determining whether a motion has been detected comprises a means for determining whether the number of pixels determined as motion pixels is equal to or greater than a predetermined value and determining that a motion has been detected when it is determined that the number of motion pixels is equal to or greater than the predetermined value.

34. The apparatus of claim 31, further comprising a means for performing a Gaussian filtering on the received image frames to smooth the image frames and transmitting the smoothed image frames to the motion detection means.

35. The apparatus of claim 21, wherein the motion detection means comprises a means for interrupting the motion detection with a predetermined period and then performing a face detection.

36. An apparatus of detecting a motion from an image provided by a digital video camera, the apparatus comprising:

a means for receiving a predetermined number of image frames;

a means for detecting a motion in the image frames using the temporal edge detection algorithm and variance of pixel values with respect to time;

a means for determining whether a motion has been detected; and

a means for storing the image when it is determined that a motion has been detected.

37. The method of claim 36, wherein the means for detecting a motion in the photographed image comprises:

a means for detecting pixels by finding zero crossings of a Laplacian with respect to time using the predetermined number of image frames;

a means for calculating the variance of pixel values with respect to time using the predetermined number of image frames; and

a means for determining whether the variance calculated for each zero crossing pixel of the Laplacian is equal to or greater than a predetermined value, and determining the pixel as a motion pixel when it is determined that the variance is equal to or greater than the predetermined value.

38. The apparatus of claim 37, wherein the means for detecting pixels by finding zero crossings of the Laplacian with respect to pixel coordinates and time comprises:

a means for multiplying each of $f(t_1)$ through $f(t_{2m-1})$ in a group of image frames by a predetermined weight and then averaging the results of multiplication to obtain Laplacian of $f(t_m)$ with respect to time, i.e., $\nabla^2 f(t_m)$;

a means for multiplying each of $f(t_2)$ through $f(t_{2m})$ in another group of image frames by a predetermined weight and then averaging the results of multiplication to obtain Laplacian of $f(t_{m+1})$ with respect to time, i.e., $\nabla^2 f(t_{m+1})$; and

a means for determining a pixel of interest as a zero crossing pixel when $\nabla^2 f(t_m)$ is negative and $\nabla^2 f(t_{m+1})$ is positive or when $\nabla^2 f(t_m)$ is positive and $\nabla^2 f(t_{m+1})$ is negative,

wherein $2m$ image frames are received, "n" pixels exist in each image frame, $f(t_1)$ through $f(t_{2m})$ denote pixel values of the pixel of interest under processing, and operations of the above three means are repeated "n" times.

39. The apparatus of claim 37, wherein the means for determining whether a motion has been detected comprises a means for determining whether the number of pixels determined as motion pixels is equal to or greater than a predetermined value and determining that a motion has been detected when it is determined that the number of motion pixels is equal to or greater than the predetermined value.

40. The apparatus of claim 37, further comprising a means for performing a Gaussian filtering on the received image frames to smooth the image frames and transmitting the smoothed image frames to the means for detecting a motion in the photographed image.